09/00518

PCT/GB98/00518

The Patent Office

The Patent Office
Cardiff Road
Newport
Gwent
NP9 1RH

5

REC'D 07 APR 1998
WIPO PCT

PRIORITY DOCUMENT

I, the undersigned, being an officer duly authorised in accordance with Section 74(1) and (4) of the Deregulation and Contracting Out Act 1994, to sign and issue certificates on behalf of the Comptroller-General, hereby certify that annexed hereto is a true copy of the documents as originally filed in connection with the patent application identified therein.

In accordance with the Patents (Companies Re-registration) Rules 1982, if a company named in this certificate and any accompanying documents has re-registered under the Companies Act 1980 with the same name as that with which it was registered immediately before re-registration save for the substitution as, or inclusion as, the last part of the name of the words "public limited company" or their equivalents in Welsh, references to the name of the company in this certificate and any accompanying documents shall be treated as references to the name with which it is so re-registered.

In accordance with the rules, the words "public limited company" may be replaced by p.l.c., plc, P.L.C. or PLC.

Re-registration under the Companies Act does not constitute a new legal entity but merely subjects the company to certain additional company law rules.

Signed

Dated 13 · 3 · 1998

Patents Form 1/77

Patents Act 1977
(Rule 16)

The Patent Office

18 FEB 1997
RECEIVED BY HAND

19FEB97 E254873-8 D00027
P01/7700 25.00

Request for grant of a patent

(See the notes on the back of this form. You can also get an explanatory leaflet from the Patent Office to help you fill in this form)

The Patent Office
Cardiff Road
Newport
Gwent NP9 1RH

1. Your reference: 44.42.63555/001

Patent application number
(The Patent Office will fill in this part): 18 FEB 1997 **9703369.0**

3. Full name, address and postcode of the or of each applicant *(underline all surnames)*:
Lindqvist, Bjorn H.
Biotechnology Centre of Oslo
University of Oslo
Gaustadalléen 21
N-0371 Oslo
Norway

Patents ADP number *(if you know it)*: 09195750

If the applicant is a corporate body, give country/state of incorporation:

4. Title of the invention: Process

5. Name of your agent *(if you have one)*:
Frank B. Dehn & Co.

Address for service in the United Kingdom to which all correspondence should be sent *(including the postcode)*:
179 Queen Victoria Street
London
EC4V 4EL

Patents ADP number *(if you know it)*: 166001

6. If you are declaring priority from one or more earlier patent applications, give the country and the date of filing of the or of each of these earlier applications and *(if you know it)* the or each application number

Country	Priority application number *(if you know it)*	Date of filing *(day / month / year)*

7. If this application is divided or otherwise derived from an earlier UK application, give the number and the filing date of the earlier application

Number of earlier application	Date of filing *(day / month / year)*

8. Is a statement of inventorship and of right to grant of a patent required in support of this request? *(Answer 'Yes' if:*
a) any applicant named in part 3 is not an inventor, or
b) there is an inventor who is not named as an applicant, or
c) any named applicant is a corporate body.
See note (d))

NO

9. **Enter the number of sheets for any of the following items you are filing with this form. Do not count copies of the same document**

Continuation sheets of this form	-
Description	23
Claim(s)	-
Abstract	-
Drawing(s)	-

10. **If you are also filing any of the following, state how many against each item.**

Priority documents	-
Translations of priority documents	-
Statement of inventorship and right to grant of a patent *(Patents Form 7/77)*	-
Request for preliminary examination and search *(Patents Form 9/77)*	-
Request for substantive examination *(Patents Form 10/77)*	-
Any other documents *(please specify)*	-

11. **I/We request the grant of a patent on the basis of this application.**

Signature Julian Cockbain **Date** 18 February 1997

12. **Name and daytime telephone number of person to contact in the United Kingdom**

Julian Cockbain
0171 206 0600

Warning

After an application for a patent has been filed, the Comptroller of the Patent Office will consider whether publication or communication of the invention should be prohibited or restricted under Section 22 of the Patents Act 1977. You will be informed if it is necessary to prohibit or restrict your invention in this way. Furthermore, if you live in the United Kingdom, Section 23 of the Patents Act 1977 stops you from applying for a patent abroad without first getting written permission from the Patent Office unless an application has been filed at least 6 weeks beforehand in the United Kingdom for a patent for the same invention and either no direction prohibiting publication or communication has been given, or any such direction has been revoked.

Notes

a) If you need help to fill in this form or you have any questions, please contact the Patent Office on 0645 500505.
b) Write your answers in capital letters using black ink or you may type them.
c) If there is not enough space for all the relevant details on any part of this form, please continue on a separate sheet of paper and write "see continuation sheet" in the relevant part(s) of the form. Any continuation sheet should be attached to this form.
d) If you have answered 'Yes', Patents Form 7/77 will need to be filed.
e) Once you have filled in the form you must remember to sign and date it.
f) For details of the fee and ways to pay please contact the Patent Office.

63555.586

Process

The present invention relates to methods for producing an in vitro peptide expression library which displays a diverse population of peptides, the expression library thus produced and use of the library to identify specific target-binding peptides. The invention also relates to specific DNA sequences which include the coding region for the peptides and which bind specifically to their translation product.

In the same way that libraries provide the reader with a vast collection of a variety of books which are retrievable, so too a molecular library provides a reference bank of molecules which may be selected and retrieved. Such libraries may contain genetic material, for example fragments of DNA sequences in a plasmid or bacteriophage, or express peptides or proteins encoded by the genetic material in the library. In the latter case to allow selection of the relevant member of the library, the expressed peptide or protein must necessarily associate with the genetic material which encodes it. Currently, this is achieved by using in vivo display of peptides on genetic packages such as cells, viruses and spores, which display individual peptides on their outer surfaces as fused parts of a display protein. The invariant moiety of the display protein in a particular library is selected to have the characteristic that it is expressed on the surface of the genetic package, for example a cell or virion and is stably associated with the cell or virion such that genetic packages expressing the target protein or peptide may be retrieved.

At present only in vivo display libraries have been described. Smith and Scott (Smith (1985), Science, 228, p1315-6; Scott and Smith (1990), Science, 249, p 386-390) describe the use of bacteriophage Fd as a display vector

for a random sequence of peptides exposed on the virion surface. US-A-5223409 of Ladner expresses families of potential binding domains on the outer surface of bacterial cells or bacteriophages. Other workers in many laboratories have similarly used such genetic packages for generating expression libraries. A lot of this work has been performed on filamentous phages like M13 which has proven to be a robust and relatively easy system to handle. However, this technology still suffers from certain drawbacks, like the time and effort required to make a library that is large enough to produce enough variants for selection. Additionally, the genetic packages used thus far must be maintained in a viable state to allow both expression of the encoded protein or peptide and also propagation of the genetic package during successive screening steps. The time involved in the propagation of selected genetic packages during the screening protocol also presents a significant time burden to the researcher. Furthermore, it is necessary in the currently used <u>in vivo</u> peptide display library to transfect the genetic material of the library into a host to allow replication and expression.

It has now surprisingly been found that a peptide expression library may be generated in which the specific translation products of the genetic material in the library are directly and covalently attached to the encoding DNA sequence in the absence of an intervening molecular structure such as an encapsulating membrane. This then obviates the use of cellular genetic packages with their inherent limitations. This advance allows rapid screening for desired peptides with cycles of selection, DNA amplification and expression. Whilst DNA amplification may involve plasmid self-replication, this may instead conveniently and rapidly be performed using standard amplification techniques, for example polymerase chain reaction (PCR).

This cell-free system is made possible by the

inclusion of a sequence within the genetic material which encodes a protein which binds covalently to its own coding sequence and which includes, or is overlapping or adjacent to, the coding sequence for the peptide for display. It will be appreciated that such binding will only be possible if the genetic material and its translation product are accessible to one another. Thus, the genetic material should preferably be devoid of sequences which effectively code for peptides or proteins which would interfere with the protein:DNA interaction. Such sequences may however be present if they are silent and not expressed, for example lacking an initiation codon or requiring an initiator species which is neither present nor generated in the library environment, or are modified such that the translation product is disabled and would not prevent the protein:DNA interaction or are present or expressed in only a minority of the members of the expression library.

Thus viewed from one aspect, the present invention provides a method of producing a peptide expression library which displays a diverse population of peptides, wherein the peptides are specifically associated with the DNA encoding them through covalent protein:DNA binding, said method comprising at least the following steps:
1) preparing an amplifiable genetic library of vector DNAs which contain a sequence encoding an amino acid sequence which binds specifically to said encoding sequence, a sequence encoding an amino acid sequence for display, but lack a sequence which in the library environment would allow the expression of peptides or proteins which would prevent said covalent protein:DNA binding and resultant peptide display, eg. by generating encapsulating membrane-forming proteins and
2) expressing the genetic library thus formed.

Proteins which interact with the DNA sequence which encodes them to initiate replication are known in the literature as *cis*-acting proteins and establish a

covalent linkage to their own DNA template. This rolling circle type of replication is commonly used among circular replicons of different origins, for example single-stranded (ss) and double-stranded (ds) DNA phages (Van Mansfield et al. (1984), Adv. Exp. Med. Biol., 179, p221-230; Baas & Jansz (1988), Cur. Topics Microbiol. Immunol., 136, p31-70), ssDNA plasmids (Gruss & Ehrlich (1989), Microbiol. Rev., 53, p231-241; Novick (1989), Ann. Rev. Microbiol., 43, p537-565), ssDNA plant viruses (Stenger et al. (1991), PNAS, 88, p8029-8033; Saunders et al. (1991), Nucl. Acids Res., 19, p2325-2330), ss and ds DNA animal viruses (Berns (1990), Microbiol. Rev., 54, p316-329; Dasgupta et al. (1992), J. Mol. Biol., 228, p1-6). In the systems studied, the initiation proteins possess a nicking-closing and topoisomerase-like activity. The best studied system is that of the ssDNA phage ϕX174, where the A protein nicks the *ori* site in the viral strand of the replicative form and forms a covalent link to the 5' end of the cleaved strand. The 3' end is thereafter extended by the host polymerase displacing the 5' viral strand and after one round of replication the parental viral strand is religated and the A protein is transferred to the progeny strand to initiate a new round of replication (Baas & Jansz, 1988, supra). The P2 A protein has also been found to cleave the *ori* site in the coding region of the A gene at a site which is devoid of secondary structure and bind to the 5' end of the cleaved strand (Liu & Haggård-Ljungquist (1994), Nucl. Acids Res., 22, p5204-5210).

Most modifications of the DNA sequence would result in a modified translation product, excluding of course minor changes which would encode the same translation product due to the degeneracy of the genetic code. This discussion is however only concerned with those DNA sequence modifications which would result in a modified amino acid sequence in the translation product. As a result of DNA sequence modification, two different

outcomes are possible for each variant. Either, the binding region (of either the DNA or protein molecule) would be affected such that binding was prevented or impaired, or no effect to binding would occur. Hence, it would be feasible to contend that if two different modifications were made to the sequence, the separate pairs of DNA:translation product would either retain or lose their binding, and that if binding was retained, the translation product (or the DNA) would bind to any DNA sequence (or protein) with the corresponding binding region. However, in contrast to what might have been expected, modification of the encoding sequence did not have this effect.

It has been found that the translation products of modified gene sequences bind specifically to the modified sequence encoding them. Thus, the creation of a multitude of different translation products which bind to their specific encoding genetic material may be realized. This finding has been used for the development of the peptide expression library described herein. This library differs from previous libraries using cells or unicellular organisms to express the peptides in several fundamental ways. The library is prepared <u>in vitro</u>, the peptide for display is presented directly on the genetic material encoding it and not on the surface of a membrane or cell wall, monovalent display may generally be achieved and this method should allow the expression of extremely high library diversities. Additionally, when PCR amplification of the genetic material of the conjugate which binds to a target is to be performed, this may be executed <u>in situ</u> on the DNA of the member of the peptide library bound to the target, as the DNA is freely accessible for binding appropriate primers and does not require prior extraction or elution from the target or non-genetic portions of the peptide library conjugate. Thus, the harsh treatment e.g. low pH, usually required for elution of the genetic material from

target-binding cells or virions prior to amplification is not necessary.

Conveniently, this invention may be performed using the phage P2 DNA replication initiation system as the basis for the in vitro peptide expression library. The P2 A protein recognizes a defined initiator sequence located within the P2 A gene on the very same DNA molecule which codes for it (*cis*-action) and specifically nicks one of the strands while forming a covalent bond with one of the free end bases at the nick site (Liu & Haggård-Ljungquist, 1994, supra). Such a protein-DNA complex constitutes a genetic conjugate which can be used for peptide display purposes. The sequence of the P2 A gene has been reported (Liu et al. (1993), J. Mol. Biol., 231, p361-374).

It is known that the P2 A protein can tolerate amino acid alterations, for example, ten histidines have been inserted in its N-terminal part without loss of function. The property of *cis*-action of A allows peptide library constructions in vitro by subjecting a library of vector DNA templates (with sequence encoding various hybrid A peptides for display and with an appropriate promoter for transcribing the A gene) to a cell-free coupled transcription/translation step. This results in hybrid A peptides binding covalently to their own template DNA. The hybrid A:DNA conjugates constitute an in vitro peptide library displaying the different hybrid A peptides which can be subjected to panning against a target. The specific hybrid A:DNA conjugates which bind to the target may be recovered, where necessary, and the genetic material may then be amplified, by for example PCR, and subjected to a coupled transcription/translation step in a cell-free extract. This cycle may then be repeated as desired to obtain an individual hybrid A:DNA clone. This may be monitored by DNA sequencing until a single DNA sequence is obtained.

As used herein in reference to the peptide

expression library which displays a diverse population of peptides, the term "peptide" is intended to cover an amino acid sequence which contains at least a sequence (which may be part or all of the sequence) which is responsible for association of the peptide with the encoding DNA (the binding moiety) and a display sequence (which may be contained within, overlap with, or be distinct from the binding sequence) which is varied in different members of the library and which may be selected through appropriate selection procedures (the display moiety).

The DNA sequences encoding the peptides for expression in libraries of the invention, comprising the display and binding moieties, wherein the nucleic acid molecules include molecules with degenerate and/or functionally equivalent sequences, form a further aspect of the invention.

The invention is conveniently performed using at least a part of the P2 A protein as the binding moiety of the peptide of the peptide expression library. Contained within, or overlapping the DNA encoding this binding moiety, or adjacent to this moiety without the presence of a termination codon between such separate sequences, is the sequence encoding the display moiety which is varied in different members of the library. Preferably the sequence encoding the display moiety is inserted in, overlaps with, or is adjacent to the sequence encoding the N-terminal of the P2 A protein.

The molecular events which allow the binding of specific translation products to their own unique encoding sequence is not known. Whilst not wishing to be bound by theory, two different mechanisms may be possible. Firstly, variations in the sequence of the DNA may result in an altered binding domain which is only uniquely recognized by the translation product thereof. Alternatively, and more likely, during the process of transcription and subsequently translation (which may

take place simultaneously), first the mRNA and then the translation product may remain associated with, or in sufficiently close proximity to, the encoding DNA such that the translation product binds to the encoding DNA through binding regions which are invariant and not dependent on modifications in the DNA sequence. In the latter case, this will allow the generation of libraries which display peptides of various sizes or protein domains attached to an appropriate binding moiety.

As used herein, the term "binds specifically to said encoding sequence" is thus intended to indicate that the amino acid sequence whilst it may not uniquely recognize its encoding DNA if isolated and introduced to a series of different DNA sequences, will bind to its own encoding sequence when produced from its encoding DNA by transcription and translation without interference.

The vector DNAs used to generate the library may be provided with means for both amplification and transcription. Suitable vector DNAs with means for amplification include double-stranded DNA with a replication origin, for example self-replicating plasmids. Alternatively, the DNA may be amplified through technical intervention, for example by providing the DNA with appropriate sites for binding primers for a amplification reaction, for example PCR. Clearly such sites would in most cases inherently be present in any vector DNA such that the appropriate choice of primers would facilitate amplification. Means for transcription include the provision of a promoter sequence and a termination sequence.

Vector DNAs must also necessarily contain diverse display peptide encoding sequences to obtain a library of different peptides for display. Such different sequences may be introduced for by example randomization as described in the literature using randomized primer sequences in PCR (Schmidt and Skerra (1993), Protein Engineering, 6, p109-122). Randomized primer sequences

may be produced using standard chemical synthesis with commercial DNA synthesizers or may be purchased commercially.

The vector DNA with the features necessary for the generation of a library forms a further aspect of the invention.

Expression of the genetic material of the library may be performed by the provision of appropriate factors necessary for transcription and translation. This may conveniently be provided in cell-free extracts from prokaryotes or eukaryotes, for example of E. coli (Nevin & Pratt (1991), FEBS, 291, p259-263). Prokaryotic (e.g. E. coli) and eukaryotic (e.g. wheat germ or reticulocyte) cell-free extracts are available commercially (Amersham/Promega). The use of eukaryotic cell-free extracts to obtain libraries of the invention, allows the preparation of translation products which may be subject to post-translational modification, for example glycosylation, phosphorylation or base derivatization. This thus provides libraries which express peptides with different characteristics not simply based on the primary sequence of the display peptide and more appropriately mimics naturally occurring moieties.

Viewed from a further aspect, the invention provides an in vitro peptide expression library which displays a diverse population of peptides, wherein the peptides are specifically associated with the DNA encoding them through covalent protein:DNA binding, and wherein said encoding sequence is carried on a vector DNA which contains a sequence encoding an amino acid sequence which binds specifically to said encoding sequence, a sequence encoding an amino acid sequence for display, but lacks a sequence which would allow the expression of peptides or proteins which would prevent said covalent protein:DNA binding and resultant peptide display.

A library generated as described herein may be used for any of the applications for which conventional in

vivo display libraries of the art are used. Such uses are well documented in the literature. For example, the library of the invention may be used to identify a peptide which binds specifically to a target molecule.

It is known in the art that peptides of different size may be arranged in an appropriate tertiary structure to produce a domain with particular steric and charge characteristics. Such a domain may, by virtue of its specific tertiary arrangement, specifically recognize or bind to a particular target molecule. Examples of such peptides include, but are not limited to binding regions of proteins and the variable binding regions of antibodies. Co-pending application (United Kingdom Patent Application No. 9525066.8) describes that much smaller peptides without defined tertiary structure may also have specific target binding properties. The peptides for display by the library of the invention may thus be small peptides, for example up to 40 amino acid residues, e.g.. 5 to 30, preferably 7 to 20 and most preferably 10 to 15 amino acid residues, which do not have a fixed tertiary structure, or may be larger peptides which form a fixed tertiary structure.

Target molecules may include small chemical compounds, for example heterocycles or pharmaceutical compounds, polypeptides, proteins, polynucleotides or any entity having distinctive surface characteristics which may be specifically recognized. Thus, for example, specific target-binding peptides may be identified which would have utility in diagnostic assays, for example in clinical procedures to assess the levels of biological or non-biological molecules in the human body or samples, extracts or material derived therefrom, or in assays which ascertain the levels of biological or non-biological materials in other non-biologically derived materials.

Target molecules, preferably in purified form, may be used to select a specific target-binding peptide-

bearing genetic conjugate from the library in a number of different ways. Conveniently, the target may be attached to a solid support and used as an affinity matrix. Numerous solid supports and methods for the attachment of molecules directly or indirectly, covalently or non-covalently (e.g. by a streptavidin-biotin or IgG-protein A coupling) are well known in the art and widely described in the literature. Thus for example, supports in the form of microtitre wells, tubes, dipstick, particles, fibres or capillaries may be used. Advantageously, the support may comprise magnetic particles e.g. the superparamagnetic bead produced by Dynal AS (Oslo, Norway and sold under the trademark DYNABEADS).

For selection, the expression library may be contacted with the target attached to a solid support. The support may be washed to remove members of the library which do not bind to the target or extracted from the expression library as appropriate for the support being used. Selected peptide:DNA conjugates may then be released from the solid support, if necessary, through disruption of the binding between the target molecules and solid support or target molecules and peptide:DNA conjugates for subsequent amplification or isolation of the genetic material. Alternatively, amplification may be performed <u>in situ</u> without disruption of the target to peptide:DNA conjugate bond or release of the genetic material from the conjugate. The target molecule may also be used as a free agent in the absence of a support. Selection may then be performed by removal of non-bound conjugates, for example by using antibodies directed to a region of the expressed peptide which is present on all members of the library and which is only accessible when not bound to target molecules. Target molecules may alternatively be provided with a means for immobilization such that this may be used to remove the target and bound peptide:DNA conjugates after mixing of the target and

library. Such means for immobilization may for example constitute one partner of a coupling pair e.g streptavidin-biotin, attached to the target molecule and the other partner attached to a support to be used for retrieval.

Thus viewed from a yet further aspect, the invention provides a method of identifying a specific target-binding peptide, said method comprising a) screening a library of the invention with target molecules and b) selecting a DNA vector expressing a peptide which binds specifically to said target molecules.

More than one cycle of screening and selection may be necessary to obtain a target-binding peptide of the desired specificity.

A selected peptide attached to the DNA vector may be isolated by separation from the genetic material, may be synthesized by transcription and translation of the genetic material which may be amplified, or may be synthesized chemically after sequencing of the appropriate DNA sequence encoding it or direct sequencing of the peptide. Chemical synthesis of the peptide may be performed by methods well known in the art involving cyclic sets of reactions of selective deprotection of the functional groups of a terminal amino acid and coupling of selectively protected amino acid residues, followed finally by complete deprotection of all functional groups. Synthesis may be performed in solution or on a solid support using suitable solid phases known in the art.

Preferably, if the affinity of the selected peptide for the target molecule is not significantly affected, only the display moiety of the peptide may be synthesized.

Appropriate target-binding peptide:DNA conjugates may be provided with a reporter molecule for use in qualitative or quantitative assays for determining the presence or absence of target molecules.

Thus, viewed from a yet still further aspect, the invention provides a method of assaying for the presence of a target molecules in a sample, said method comprising (a) contacting said sample (e.g. of biological, biologically-derived or non-biological material) with a molecular probe comprising (i) a peptide moiety capable of selectively binding to said target molecule selected from the library of the invention and (ii) a reporter moiety; and (b) directly or indirectly assessing the target bound probe.

Bifunctional molecular probes (comprising (i) and (ii) as described above) for use in the assay form a further aspect of the invention.

In this assay method, assessment of the binding of the bifunctional compound to any of the targets to which it is specific, that is present in the sample, may be direct or indirect. Direct and indirect assessment are well known in the field of diagnostic assays. Such procedures may involve separation of the bound (or unbound) bifunctional compound either of which may serve as the analyte. Assessment of the target molecule:bifunctional compound conjugate may be qualitative or, more preferably, quantitative and will involve direct or indirect assessment of the reporter moiety.

The assay may be directed to the assessment of a second target with the first target, in which the reporter moiety on a probe for the second target is recognized by the bifunctional compound. Thus a bifunctional compound may be directed to a probe, preferably molecular, which recognizes a further target, in which case the probe is allowed to bind to the further target under suitable binding conditions prior to the addition of the bifunctional compound as mentioned above.

To provide the probe, the specific binding peptide:DNA conjugates may incorporate or be conjugated to a reporter moiety such that the presence within a test

sample of the target of interest may be determined and/or quantified.

The peptide moiety in the bifunctional compound will derive its selective binding capability from the presence of a binding moiety (distinct from the binding moiety which recognizes DNA) which constitutes some or all of the amino acid residues of the peptide.

The reporter moiety may be any moiety capable of direct or indirect detection, e.g. by virtue of its enzymatic properties, radiation emission, scattering or absorption properties, of its magnetic properties, or of its ability to cooperate with or bind to a complimentary agent to produce a detectable effect e.g. interact with an enzyme to produce a signal, gas evolution, light emission, colour change, turbidity, precipitation, etc. The reporter moiety may alternatively be any part of the peptide:DNA conjugate which is recognizable and may bind a further molecule which may directly or indirectly produce a signal. Thus, for example, an antibody directed to a particular region of the genetic material or peptide may be employed. The above-mentioned moieties are well known within the field of diagnostic assays.

The reporter moiety in the bifunctional compounds of the invention may be incorporated in or conjugated to the peptide or DNA moiety. Thus by way of example radiolabelled amino acids or nucleotides may be used for the construction of the peptide or encoding DNA, the radionuclides built into the peptide or nucleic acid structures then functioning as the reporter moieties. Such labelled constituents may be incorporated during the preparation of the parent libary or during subsequent screening or amplification steps where these are performed. Alternatively a reporter molecule may be conjugated to the peptide or DNA which directly or indirectly allows detection or measurement of the presence of the target to which the peptide is capable of binding. Such reporter molecules include for example

radiolabels, chemical labels, for example chromophores or fluorophores (e.g. dyes such as fluorescein and rhodamine), or reagents of high electron density such as ferritin, haemocyanin or colloidal gold. Alternatively, the reporter molecule may be an enzyme, for example peroxidase or alkaline phosphatase, wherein the presence of the enzyme is visualized by its interaction with a suitable entity, for example a substrate. Coupling of enzymes to peptides may be achieved using conventional techniques, e.g. using an activated enzyme such as activated alkaline phosphatase (Boehringer Mannheim Biochemicals). The reporter moiety may also form part of a signalling pair wherein the other member of the pair is found on, or in close proximity to, the target to which the peptide binds, for example, a fluorescent compound and a quench fluorescent substrate. As mentioned previously, the peptide or DNA may also be detected by association with, or binding of, a further molecule which recognizes its identity, for example an antibody directed to part of the sequence which may form the target-binding moiety of the peptide or a region of the peptide not involved in target binding which may be added for the purposes of recognition, or in the case of DNA directed against specific nucleic acid motifs. Thus the specific binding peptide may be a region within a larger peptide, wherein the portions of the peptide not involved in binding the target may function as a reporter moiety, or as a linking group linking the specific target-binding region to a reporter moiety or to a further component of the probe, e.g. a carrier or a macromolecule.

The bifunctional compounds useful in accordance with the invention can be produced by conjugating a reporter molecule to the resulting appropriate peptide, either directly or via a linker moiety. Generally this will be by reaction with an optionally activated carboxyl or amine functionality on the peptide. Such conjugation reactions are well within the ability of a chemist of

ordinary skill.

Alternatively, the reporter molecule may be introduced by utilizing an appropriately labelled amino acid in the construction of the peptide.

The bifunctional probe compounds may be used to recognize specific targets of interest in various systems known in the art, including diagnostic assays as mentioned previously.

Libraries according to the invention also have utility in screening protocols for identifying compounds with appropriate biochemical, biological or structural properties, for example to identify peptides which have certain biochemical activity in a defined assay. Such peptides may be used for the preparation of compounds with the particular activity for example, inhibitors, activators, or catalysts of certain reactions or interactions.

By this method, peptides with enzymatic, inhibitory or stimulating properties may be identified which may have utility in for example the pharmaceutical field.

The following Examples are given by way of illustration only.

Example 1:

General methodology for generating an in vitro peptide library and panning for a target

Materials:

A. Plasmid or PCR fragment containing T7 promoter, ribosome binding site, the P2 A gene and T7 terminator. Such plasmids have been described by Liu & Haggård-Ljungquist (1994, supra) or may be obtained from Bionor Research Products AS (BRP) or Biotechnology Centre of Oslo, University of Oslo (BiO).

B. One primer (library primer) that contains the following sequences complementary to plasmid/fragment:
   - T7 promoter,
   - ribosome binding site
   - 30 random nucleotides (XXT/G) after the first ATG start codon, alternatively, one cysteine codon after the first start codon and after the random sequence (for constrained peptide libraries).
   - approximately 20 nucleotides downstream from the first start codon complementary to the coding sequence for the P2 A gene.

Primers may be custom synthesized or obtained from BiO.

C. One PCR primer T7 promoter region (BiO).

D. One PCR primer in the T7 terminator region (counter clockwise) (BiO).

E. Target bound to solid support. Conveniently this may be performed using a biotinylated target and binding this to streptavidin or avidin on a solid support. Alternatively, avidin itself may be the "target" if avidin-binding peptides are sought. Streptavidin bound to microtiter wells or streptavidin bound to magnetic particles or Avidin-Resin may be bought from Dynal (Norway) or Promega (USA), respectively.

F. T7 S30 extract for <u>in vitro</u> coupled transcription/ translation of linear templates may be obtained from Promega (USA).

* The rest of the material needed is standard for anyone working with molecular biology techniques.

Methods:

1. Starting with a plasmid or PCR fragment as mentioned in A, linear PCR is performed by adding the library primer mentioned in B. The exact set-up for this reaction is dependent on the primer and the same considerations that apply to PCR or cycle sequencing also apply here. This will generate a library of up to 10^{12} to 10^{13} molecules, dependent on the effectiveness of the PCR. To avoid primer competition in the next step, the remaining library primers should preferably be removed at this point by the use of a Centricon-100 column (Amicon).

2. To amplify this material, 5-7 cycles of PCR with primers C and D are performed. This is performed using the library primer extended DNA aliquoted into five portions.

3. One portion of the material from the library generated (ex. one fifth) is added to a S30 extract for linear fragments containing T7 RNA Polymerase (F) as described in the Promega manual. The reaction is incubated for 60 min. at 37°C and then stopped by placing the tube(s) on ice.

4. The target is attached directly to a solid support by the biotin-streptavidin system. Avidin coupled to Resin matrix or streptavidin magnetic beads may be obtained commercially from Promega and Dynal, respectively.

5. The S30 extract is diluted 1:10 in the desired binding buffer (Sambrook et. al. (1989), Molecular Cloning : A laboratory manual, 2nd Ed., Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y.), and the peptide library in the S30 extract is allowed to interact with the target for 1-3 hours (or overnight).

Non-binders are removed by washing 5x with 1XPBS +0.5% Tween-20 (Sigma) for 5 minutes. The bound peptide-protein A-DNA complex is eluted from the target with the desired eluant, for example biotin if the target is avidin and an avidin-binding peptide is sought. Elution with boiling dH_2O may also be performed to release the complex from the target and simultaneously release the genetic material from the non-genetic material.

6. The eluted DNA is purified before going into the next round of PCR. This is performed most conveniently by using a Centricon-100 column and following the manufacturer's recommendations. The final volume of the eluted DNA is 50 μl. Alternatively the complex may be purified before PCR without separation of genetic and non-genetic material.

7. A new PCR reaction is set up using primers C and D with 30-40 cycles.

8. The whole procedure from step 3 to 7 is performed again, four to five times.

9. After the final cycle of elution and PCR the fragments need to be cloned in order to isolate and study individual fragments to determine their sequences. This is performed by digesting the final PCR fragment with Xba I and BamH I and ligating this to vector pET-3a (Novagene) (A) digested with the same enzymes.

10. The ligated vector is then transformed into E.coli. Since there are many copies of the same sequence, the efficiency of transformation is not critical.

11. Individual clones are picked and the plasmid DNA isolated by standard protocols. A final round of S30 extract and selection (steps 3 to 7) is performed to

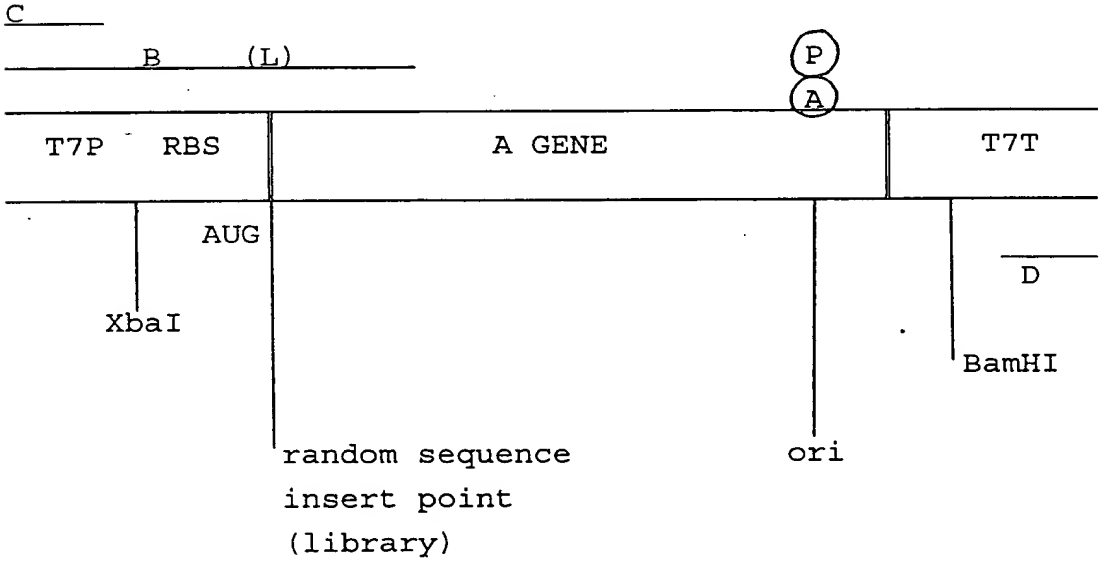
prevent binders that only act co-operatively (together with other binders).

12. The final PCR product is sequenced over the variable region, and a consensus sequence is obtained. In order to obtain a good consensus sequence, up to 50 clones should be sequenced.

13. The deduced peptide sequence is synthesised and tested separately for its binding properties.

Example 2

A library population is obtained using a randomized base primer for amplification of the A gene. The display module which corresponds to about 3000 base pairs is shown schematically as follows for the linearized plasmid pEE709 (A gene inserted into pET8c= pET3d at the NcoI site after fill in).



In which:

T7p = T7 promoter ø10
T7t = T7 terminator

RBS = ribosome binding site
AUG = start codon for the A protein
A = A protein
P = displayed peptide/protein
B = library strand primer containng random sequence (L) as defined in Example 1
C = PCR primer (T7p)
D = PCR primer (T7t)

In the above diagram the A gene insert starts with GCC (the second codon) and ends with GCA (bases 3427-29). (See DNA sequence of Liu et al., 1993, supra).

The primers used in the Example are as follows:

B: GAAATTAATACGACTCACTATAGGGAGACCACAACGGTTTCCCTCTAGAAA TAATTTTGTTTAACTTTAAGAAGGAGATACCATG - $[XXT/G]_{10}$ - GCCGTTAAA GCCTCCGGG [125 nucleotides]
C: GAAATTAATACGACTCACTATAGGG
D: CAAAAAACCCCTCAAGACCCG

1. Generation of a Peptide Library

A DNA fragment population with a set of randomized bases is obtained as follows:

Linear PCR (or primer extension) is performed on the linearized (HindIII) plasmid pEE709 by the library primer B. The reaction mixture of 100 μl contains: 0.3 μg plasmid DNA (about $6x10^{10}$ molecules or 0.1 pmol), 5 μg library primer DNA (about $7x10^{13}$ molecules or 125 pmol) and the rest of the ingredients as described for the PCR reaction below. The mixture is subjected to 5 cycles of PCR as described below. The library primer is preferably removed using a Centricon-100. The library primer extended DNA (library) is diluted and subdivided into five 100 μl (final volume) aliquots and subjected to a

limited PCR (5 cycles) using primers C and D. Each reaction mixture contains: 0.6 μg library primer extended DNA, 125 pmol of the respective primers C and D, 0.2 mM of each dNTP, 50 mM KCl, 4 mM $MgCl_2$, 10 mM Tris-HCl (pH 9.0 at 25°C), 0.1% Triton X-100 and 2.5 U Taq DNA polymerase (Promega) in a final volume of 100 μl. The mixture is subjected to 35 cycles of 1 minute at 94°C, 2 minutes at 42°C and 3 minutes at 72°C in a thermocycler (Perkin Elmer model PCR1000). The PCR product is purified by removal of the primers (Centricon-100), phenol treatment and ethanol precipitation. At this point the library should comprise 10^{12} to 10^{13} DNA molecules.

An alternative library approach would simply be to run PCR cycles on the vector fragment DNA using the library primers B and D to drive the PCR. In this case, however, sequence biases are likely to be generated and library heteroduplexes will appear during amplification.

2. <u>In vitro translation and screening for an avidin binding peptide</u>

A. A combination of Promega's T7 S30 and S30 linear template extract is used for coupled transcription/ translation of linear DNA templates. Transcription of the A gene is driven by the T7 RNA polymerase from the T7 promoter ø10. One of the five DNA library sets which is phenol treated and precipitated by ethanol is resuspended in 9 μl distilled water. To this volume are added the ingredients (5 μl of aminoacid mix, 20 μl of S30 pre-mix, 1 μl of T7 S30 and 15 μl of S30 for linear templates) of the S30 protocol (Promega) to make a final volume of 50 μl. The coupled transcription/translation process is allowed to proceed for 60 minutes (or as long as required) at 37°C.

B. The reaction mixture (50 µl) is added to 50 µl SoftLink Avidin Resin (Promega) and allowed to mix for 2 hours at room temperature for the panning of peptides binding to Avidin. The Resin is pelleted by centrifugation (10000 rev/min for 5 minutes) and washed (five times with PBS, 20 mM Na_2HPO_4, 100 mM NaCl pH 7.5). Potential avidin binders are eluted with 5 mM biotin or simply by subjecting the entire avidin-Resin complex to PCR with primers C and D as described in 1A. The PCR product is separated from the avidin-Resin by centrifugation, phenol treated and precipitated by ethanol before being subjected to a new coupled transcription/translation and panning cycle. Cycles of peptide display and panning can be repeated until the anticipated peptide enrichment has been achieved. Polyclonal antibodies specific for the A protein can be used to monitor the presence and increase of the protein A carrier during the panning. After the fourth round of panning the final PCR product is cut with restriction enzymes Xbal and BamHI and inserted into pET-3a (cut with the same enzymes) by ligation. After transformations, individual colonies are isolated and plasmids extracted for sequence determination of the insert in order to obtain the amino acid sequence of the peptide.

① 98/00 S[redacted]

② 18 2 98

③ Frank B Dehn